



# Action Categorization in Rhesus Monkeys: discrimination of grasping from non-grasping manual motor acts

## Citation

Nelissen, Koen, and Wim Vanduffel. 2017. "Action Categorization in Rhesus Monkeys: discrimination of grasping from non-grasping manual motor acts." Scientific Reports 7 (1): 15094. doi:10.1038/s41598-017-15378-6. <http://dx.doi.org/10.1038/s41598-017-15378-6>.

## Published Version

doi:10.1038/s41598-017-15378-6

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:34492871>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# SCIENTIFIC REPORTS

OPEN

## Action Categorization in Rhesus Monkeys: discrimination of grasping from non-grasping manual motor acts

Koen Nelissen<sup>1</sup> & Wim Vanduffel<sup>1,2</sup>

Received: 8 August 2017

Accepted: 25 October 2017

Published online: 08 November 2017

The ability to recognize others' actions is an important aspect of social behavior. While neurophysiological and behavioral research in monkeys has offered a better understanding of how the primate brain processes this type of information, further insight with respect to the neural correlates of action recognition requires tasks that allow recording of brain activity or perturbing brain regions while monkeys simultaneously make behavioral judgements about certain aspects of observed actions. Here we investigated whether rhesus monkeys could actively discriminate videos showing grasping or non-grasping manual motor acts in a two-alternative categorization task. After monkeys became proficient in this task, we tested their ability to generalize to a number of untrained, novel videos depicting grasps or other manual motor acts. Monkeys generalized to a wide range of novel human or conspecific grasping and non-grasping motor acts. They failed, however, for videos showing unfamiliar actions such as a non-biological effector performing a grasp, or a human hand touching an object with the back of the hand. This study shows the feasibility of training monkeys to perform active judgements about certain aspects of observed actions, instrumental for causal investigations into the neural correlates of action recognition.

The ability to recognize other individuals' actions is of significant importance for primates and a fundamental aspect of social behavior. Rhesus monkeys' ability to recognize, or even comprehend actions in terms of goals and intentions, is often assumed in many neurophysiological studies that use action observation as a proxy for action recognition or comprehension<sup>1–7</sup>. Since most of these studies did not include a behavioral response, it is difficult to quantify whether monkeys actually recognize these observed motor acts and to what degree neural activity during these observation tasks reflects action understanding. Several protocols have been used to study the monkey's action comprehension abilities in a more ethological setting. These include experiments during which monkeys observed accidental versus intentional actions, rational versus irrational actions or actions performed by actors in an unwilling versus an incapable situation. Monkey behavioral responses measured during these experiments include different paradigms such as looking-time, forced-choice food foraging, or observation of monkey's vocal and body responses towards the actor<sup>8–13</sup>. Most of these behavioral studies, however, were performed in a setting not easily compatible with simultaneous recording of brain activity.

While several action recognition models have been proposed (for recent review, see<sup>14</sup>) to date the exact mechanisms and detailed computations performed by the brain underlying action recognition, are far from understood. To gain further insights into the neuronal correlates of action recognition, tasks will be required that allow measuring functional brain responses not only when subjects passively observe actions, but when they have to make behavioral judgements related to the observed actions<sup>15–18</sup>. Combining these behavioral tasks with reversible perturbations<sup>19</sup> will provide much needed causal evidence of specific visuo-motor nodes roles in action recognition. For instance, numerous studies using categorization tasks combined with focal reversible perturbations have allowed examining the causal role of certain brain regions in the perception of specific stimulus properties or categories in monkeys<sup>20–23</sup>. In humans on the other hand, in particular TMS in combination with behavioral

<sup>1</sup>Laboratory for Neuro- & Psychophysiology, Department of Neurosciences, KU Leuven, Leuven, 3000, Belgium.

<sup>2</sup>Massachusetts General Hospital, Harvard Medical School, Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, Massachusetts, 02129, USA. Correspondence and requests for materials should be addressed to K.N. (email: [koen.nelissen@kuleuven.be](mailto:koen.nelissen@kuleuven.be))

tasks has been employed to provide causal evidence for specific visuo-motor contributions to different aspects of action recognition<sup>24–26</sup>, for review see<sup>27</sup>.

In this behavioral study, we investigated rhesus monkeys' ability to discriminate different types of motor acts. In a two-alternative forced-choice action categorization task, monkeys learned to discriminate grasping motor acts (Fig. 1a, Supplementary Video S1) from other manual non-grasping motor acts including touch with a finger (Fig. 1b, Supplementary Videos S2 and S5), touch with a closed fist (Supplementary Video S3), mimicked grasp next to an object (Supplementary Video S4). The monkeys indicated their choices by making a saccade to the left or right (Fig. 1a,b, right panels). After monkeys reached proficiency on this task (above 80% correct trials), we tested how well monkeys generalized this learned categorization rule (grasping versus non-grasping) to new, untrained videos of either grasping or non-grasping manual motor acts. During these generalization sessions, we tested if monkeys could discriminate untrained videos depicting 1) novel viewpoints of grasping and non-grasping motor acts, grasping with 2) novel objects and 3) novel effectors (human, conspecific or artificial), and non-grasping motor acts with 4) novel spatial positions of familiar hand configurations and 5) novel hand configurations.

## Methods

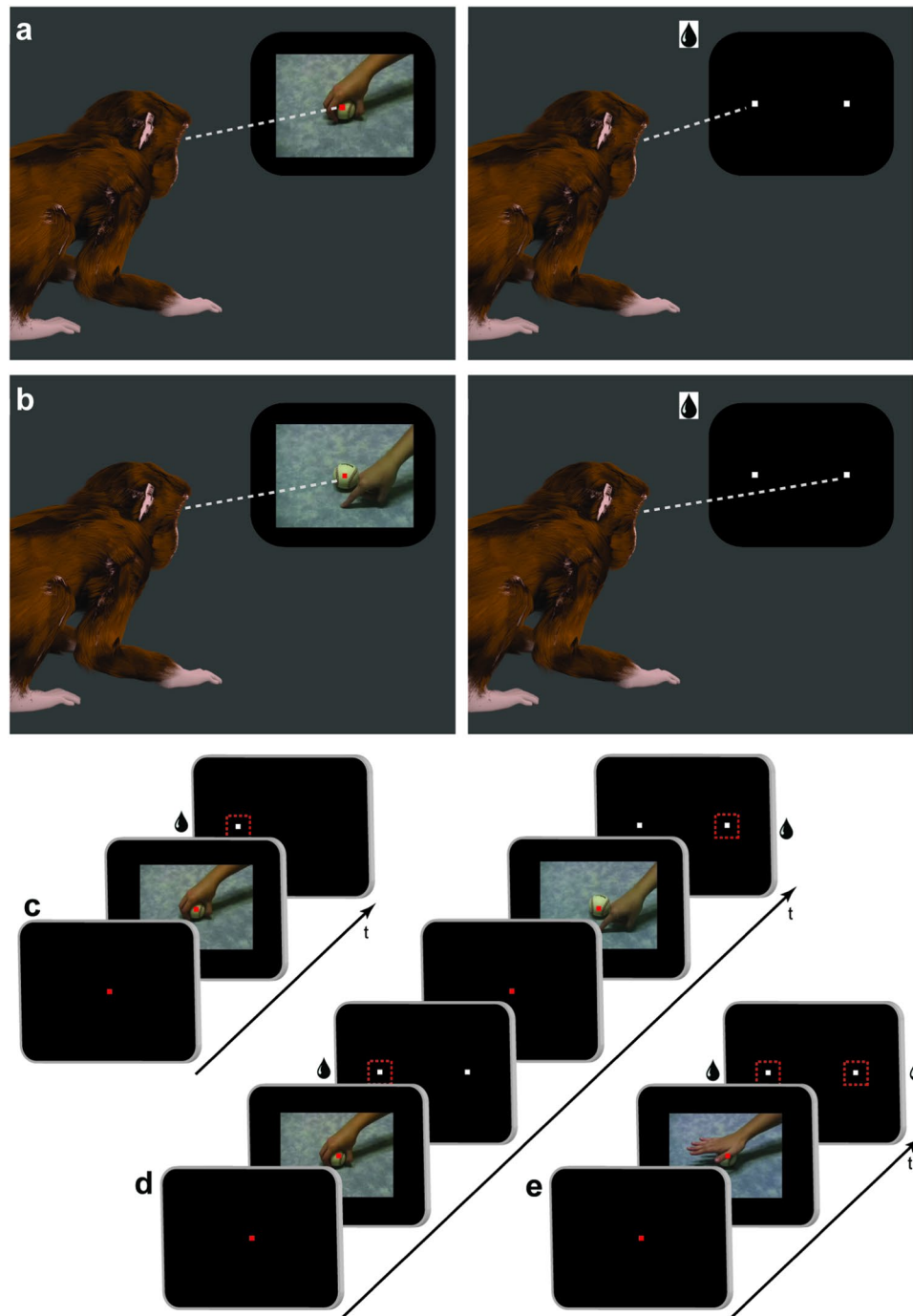
**Subjects.** Two male (M1, M2) rhesus monkeys (*Macaca mulatta*, 4–6 kg, 3–5 years old) participated in the experiments. All animal care and experimental procedures met the national and European guidelines and were approved by the animal ethical committee of the KU Leuven.

**General Fixation Training.** The monkey subjects were accustomed to sitting in a sphinx position in a plastic monkey chair, directly facing a liquid crystal display (LCD) screen (60 Hz frame rate), which was positioned at 57 cm from the monkeys' eyes<sup>28,29</sup>. During initial training, they were required to maintain fixation within a  $2^\circ \times 2^\circ$  window centered on a red fixation target (size:  $0.18^\circ \times 0.18^\circ$ ) in the middle of the screen. Eye position was monitored at 120 Hz through pupil position and corneal reflection (Iscan). During this initial training phase, the monkeys were rewarded (fruit juice) for fixating the small red target within the fixation window for long periods (up to several minutes).

**Categorization task and training.** The monkeys were trained to perform a two-alternative action categorization task, during which they had to discriminate videos of grasping motor acts from videos showing other manual non-grasping motor acts. Each categorization trial consisted of the following sequence of events: at the beginning of a trial, the monkey was required to fixate a small red fixation target (size:  $0.18^\circ \times 0.18^\circ$ ) in the center of the screen, after which a video was presented with the fixation target superimposed (Fig. 1a,b, left panels). During video presentation, the monkey had to hold fixation within a  $2^\circ \times 2^\circ$  degree window centered on the fixation target. After presentation of the video, the central fixation target and the video were replaced by two peripheral white fixation targets (size:  $0.18^\circ \times 0.18^\circ$ ), located along the horizontal meridian at  $9.25^\circ$  to the left or the right of the center of the screen. The monkey then had to make a saccade to one of the two targets in order to receive a juice reward (Fig. 1a,b, right panels). Grasping motor acts (Fig. 1a, Supplementary Video S1) were associated with the left target and non-grasping motor acts (Fig. 1b, Supplementary Videos S2–5) with the right target. Trials would be aborted if the monkey did not hold his gaze within the  $2^\circ \times 2^\circ$  central window for the entire video presentation or if monkey failed to saccade to one of the two peripheral targets within 2 seconds after stimulus presentations. These aborts were not incorporated into the data analysis.

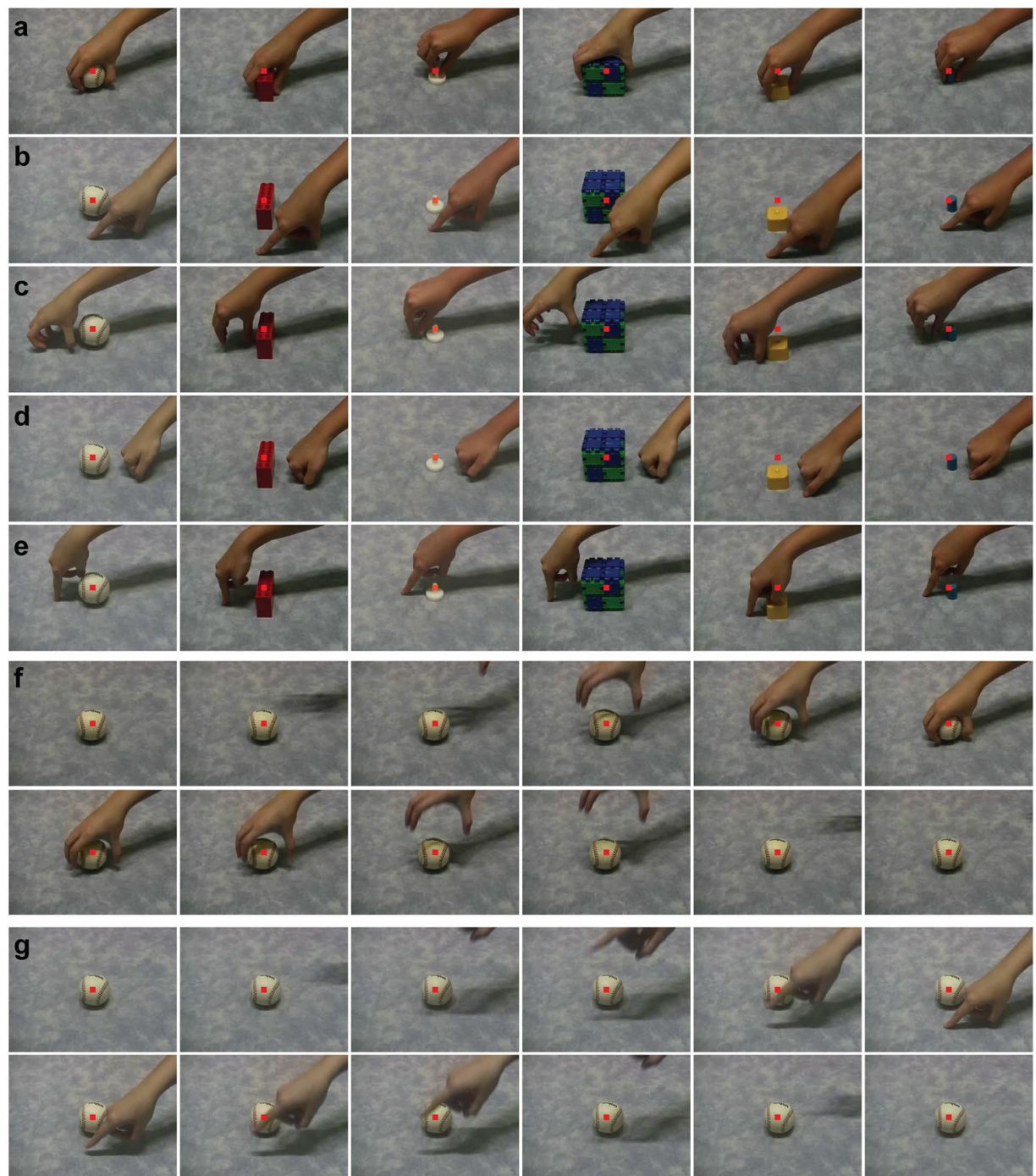
During the first training sessions, in some of the trials (up to 30% of the trials) only the correct peripheral target was presented after the video (Fig. 1c). This way monkeys learned the general procedure of initial fixation followed by a saccade to a peripheral target. Through association, monkeys learned the correct target location for the grasping versus the non-grasping videos. Initial training started with only 2 videos, depicting grasping of a baseball (Fig. 1a, left panel) or a finger touching the surface in front of the baseball (Fig. 1b, left panel). Gradually, during the following training sessions, more stimuli were added to the stimulus set, depicting grasping motor acts with the 18 different objects (6 examples shown in Fig. 2a) or non-grasping motor acts consisting of the finger touch front, finger touch side, fist touch or mimicked grasps (6 examples shown in Fig. 2b–e). During training sessions, stimuli were picked randomly from this stimulus set. During these first categorization sessions, the rest of the trials consisted of two peripheral targets, and monkeys were rewarded for choosing the correct target. During later sessions, all trials consisted of two peripheral targets (Fig. 1d). In addition, we used a response-bias-correction procedure, similar to<sup>15</sup>. During this procedure, a trial in which an error occurred was followed by a trial of the same stimulus category. This procedure was maintained until the monkey made the correct response. Both monkeys typically performed around 600–700 trials during a daily training session.

**Stimuli.** A total of 90 videos (grasping and non-grasping) were used in the training session. Grasping videos consisted of a human hand grasping diverse objects (18 different objects in total) with various grips including precision grip, 3-finger grasp, and whole hand grasp. Figure 2a shows example frames of 6 of the objects that were used. The non-grasping motor acts consisted of the same human hand touching the surface next to the same 18 objects used in the grasping videos. These non-grasping motor acts included touching the surface (table) with a finger (Fig. 2b,e), mimicking a grasp next to the object (Fig. 2c) or touching the surface next to the object with a closed fist (Fig. 2d). The grasping videos all started with a static, centrally-positioned object, followed by a human hand appearing in view and grasping the object, after which the hand released the object and disappeared out of view (Fig. 2f). A non-grasping video followed the same sequence of events: a video started with a static object, after which a hand came into view and touched the surface next to the object, and finally disappeared out of view (Fig. 2g, finger touch front). The videos measured  $13.9$  by  $10.4$  degrees, with a duration of 3 seconds.



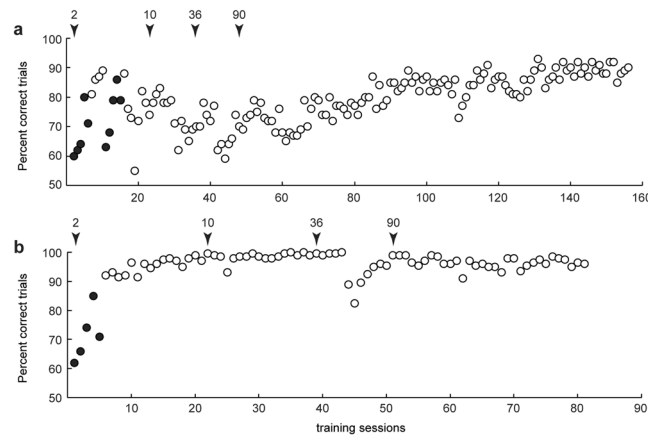
**Figure 1.** Two-alternative categorization task training and generalization testing procedure. Monkeys were required to fixate a video on a screen positioned in front of them showing either a grasping (a) or a non-grasping (b) manual motor act. After video presentation, the video disappeared and two peripheral targets appeared on the screen. A liquid reward was given if monkeys made a saccade to the correct target (left for grasping, right for non-grasping motor acts). Monkey renderings in (a) and (b) were made using open-source software Blender (<https://www.blender.org/>). (c) In the first training sessions, during a proportion of categorization trials (up to 30%) only the correct target was shown after video presentation (for illustration only a grasping trial is shown). A saccade to this target would be rewarded. The remainder of the trials consisted of two targets as shown in (d). The red dashed squares around the targets in A,B and C were not physically shown and are for illustration purpose only to indicate a saccade to that target would result in a reward. (d) During later training sessions, after video presentation, two targets were shown simultaneously and a reward was given for a saccade to the correct target. (e) During generalization testing, after video presentation, two targets would appear and selection of either target would be rewarded in order to avoid the monkey learning the novel stimuli.





**Figure 2.** Visual stimuli used in the categorization task. (a) Six examples of different objects being *grasped*. (b–e) Six examples each of the different non-grasping motor acts: finger touch front (b), mimicked grasp (c), fist touch (d), finger touch side (e). (f) Frames of a grasping motor act: grasping videos showed a human hand approaching from the right and grasping a centrally-positioned object, after which the hand released the object and moved out of the frame. Monkeys were required to fixate the central small red spot superimposed on the video. (g) Frames of a non-grasping motor act: the non-grasping videos showed a human hand approaching from the right and touching the table without interacting with the object. In this example of a finger touch motor act, a finger touched the surface in front of the object, after which the hand moved out of sight.

**Generalization testing.** After both monkeys had reached proficiency on the categorization task (consistently above 80% correct trials), we tested their abilities to generalize the learned categorization rule to novel, untrained examples of grasping and non-grasping manual motor acts. In 13 different generalization sessions, we tested if monkeys could correctly categorize untrained videos with 1) novel viewpoints, 2) novel objects, 3) novel effectors (human, conspecific or artificial), 4) novel spatial positions of familiar hand configurations and 5) novel



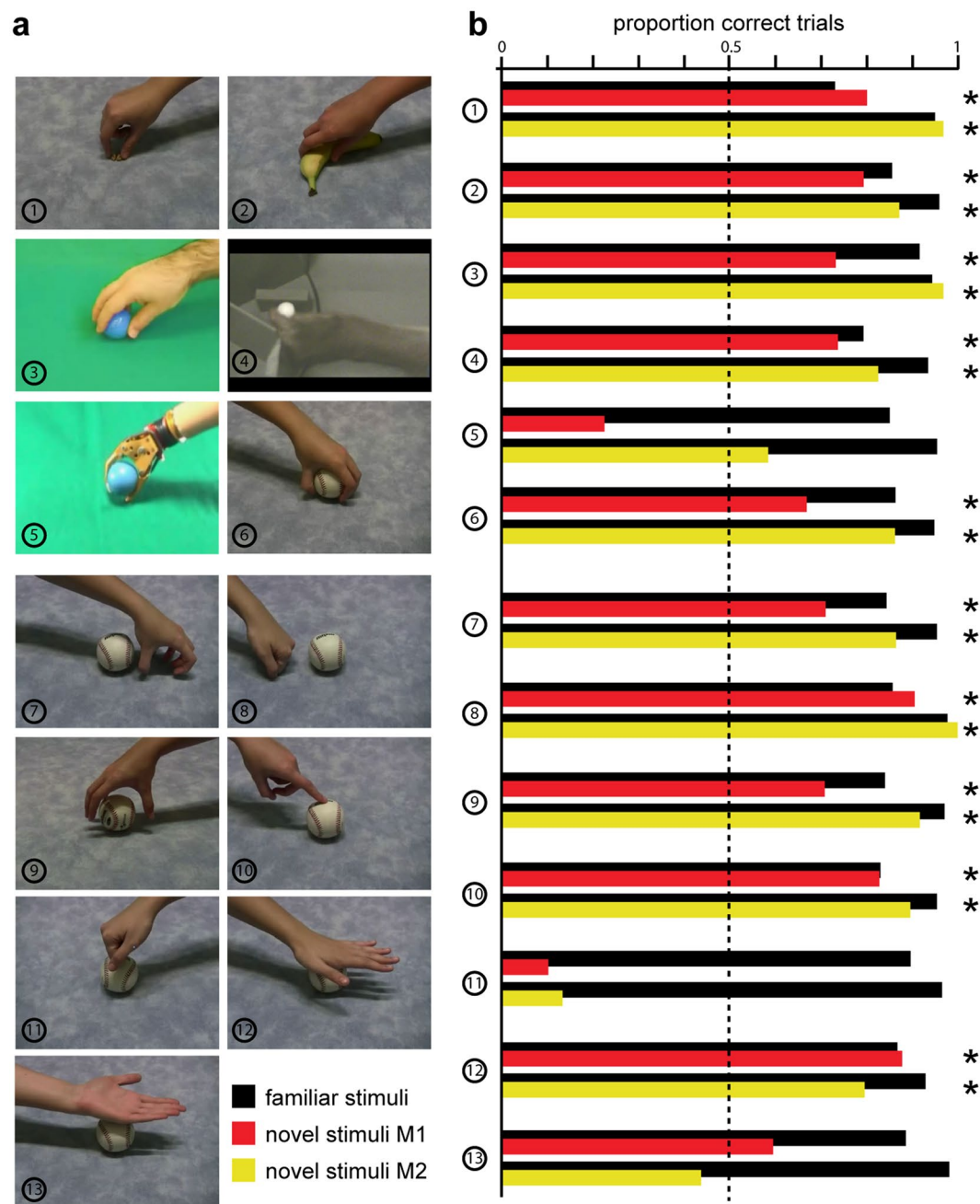
**Figure 3.** Performance as function of training session. Categorization training results of monkey M1 (a) and M2 (b). Percent correct trials are shown for different daily training sessions. Filled circles indicate session during which a proportion of trials (up to 30%) consisted of single targets (Fig. 1c). The open circles indicate sessions in which all trials consisted of two targets (Fig. 1d). A bias correction was used during training sessions only (see methods). Black arrows indicate total number of different videos in the training stimulus set.

hand configurations. Six generalization tests involved novel, untrained examples of grasping motor acts, while in the 7 additional tests, behavioral responses were tested towards novel, untrained non-grasping motor acts. During these generalization tests, around 90% of the trials consisted of trained videos (familiar) which were rewarded only on the correct side. The other 10% of the trials consisted of the novel untrained videos and were rewarded on both sides (Fig. 1e), to avoid the monkey also learning these novel stimuli<sup>15,30</sup>. No bias correction was used during these generalization sessions. During these generalization sessions, monkeys averaged 650 trials per session (which thus included ~65 trials with novel videos). For generalization tests 1 through 5, only one example of a novel motor act was used. For tests 6 through 13, the novel videos for that particular motor act consisted of the 18 different objects on which the monkeys had been previously trained (see methods).

## Results

**Categorization learning.** Figure 3 shows the training curves of both monkeys. Filled circles indicate sessions during which a proportion of trials consisted of only a single target used to shape the animal's response (see methods). Open circles indicate sessions during which all trials included both targets, in combination with a bias-correction (see methods). Initial training started with only two different videos (1 grasping, 1 non-grasping motor act) while in later phases more videos were gradually introduced. The arrows and corresponding numbers above the graphs (Fig. 3a,b) indicates the number of videos used during those time points in the course of the categorization training. While training in monkey M1 took several months (~90 sessions) before reaching a stable performance level using the 90 different videos (>80% accuracy), monkey M2 reached the same level of performance much sooner (i.e. after ~50 sessions).

**Generalization tests.** Next, we investigated whether monkeys could generalize to untrained examples of grasping and non-grasping manual motor acts. These generalization tests allowed us to examine whether the monkeys had merely memorized the familiar videos seen during the extensive training and to ascertain how specific the acquired categorization performance was for certain aspects of the videos. The results of the generalization tests are shown in Fig. 4. For generalization tests 1 through 5, only one novel video was used (frame shown in Fig. 4a). For tests 6 through 13, the novel videos for that particular motor act consisted of the 18 different objects on which the monkeys had been previously trained (see methods). As an illustration, a frame of one of these novel videos involving the baseball as the object is shown in Fig. 4a (6 to 13). Figure 4b shows the proportions of correct trials during these generalization tests for both monkeys. No bias correction was used during the entire generalization tests. Black bars indicate the performance of each monkey for the familiar action videos (90% of the trials) during the generalization tests, which is an indication of how well the monkey performed the learned classification task during that particular session. Colored bars indicate performance (proportion correct trials) for the 10% trials during which novel, untrained videos (shown in Fig. 4a) were shown (monkey M1: red bars; monkey M2: yellow bars). Asterisks indicate significant generalization (binomial test,  $p < 0.05$ ). For generalization tests 1 through 6, in which a novel example of a grasping motor act was shown, significant generalization therefore indicates that monkeys categorized these novel videos correctly as 'grasping'. For generalization tests 7 through 13 in which a novel example of a non-grasping motor act was shown, significant generalization indicates that the monkeys categorized these novel videos correctly as 'non-grasping'. As shown in Fig. 4b, monkeys generalized to untrained grasping motor acts depicting an object grasped with a precision grip (1) or whole hand grasp (2). Both monkeys could also generalize to a novel human actor performing a grasp (3) or to a conspecific grasping (4). However, both monkeys failed to generalize to grasping performed by an artificial prosthetic arm (5). Monkeys also generalized to a novel untrained viewpoint (mirror image of the trained viewpoint). This was the case for grasping motor acts (6) as well as for the non-grasping motor acts including a hand mimicking a grasp (7) or a fist touching the surface in front of the object (8).



**Figure 4.** Generalization results. **(a)** Frames of the novel videos used during the 13 different generalization test sessions. In 6 different sessions, novel untrained videos showing a grasping motor act were used. These included an object being grasped with a precision grip (1) or a whole-hand grasp (2), a male actor grasping a ball (3), a monkey grasping a small cylinder (4), a prosthetic arm grasping a ball (5) and finally, mirrored versions of the familiar grasping videos (6). In 7 additional generalization tests, we tested generalization to novel, untrained non-grasping motor acts. These included mirror versions of the familiar mimicked grasp (7) and fist touch (8) videos, a hand initiating a grasp which halted immediately prior touching the object (9), a finger touching the object (10), a fist touching the object (11), an open hand touching the object with the palm (12) and an open hand touching the object with the back of the hand (13). **(b)** Proportion correct trials during the generalization tests. Black bars indicate performance of the monkeys during the 13 generalization tests for the trained familiar videos (90% of trials). Red (monkey M1) and yellow (monkey M2) bars indicate performance from the monkeys on the novel, untrained videos (10% of trials). Asterisks indicate significant generalization ( $p < 0.05$ , binomial test). No bias correction was used during the generalization sessions.

During training, all non-grasping motor acts involved actions directed away from the object, depicting situations during which the human hand did not touch or interact with the object. Thus a possible strategy for the monkeys to solve the categorization task might have been to categorize the videos according to objects being



touched or interacted with. To investigate this possibility, we also tested the 3 familiar non-grasping motor acts (mimicked grasp, finger touch and fist touch) in a novel setting where these were directed towards the object. Videos of a hand initiating a grasp that stopped before making contact with the object were categorized correctly as a non-grasping motor act (9). While a finger touching an object was categorized correctly as a non-grasping motor act (10), both monkeys nonetheless failed to correctly categorize a closed fist touching an object and consistently categorized this motor act as a grasp (11). Note that a novel viewpoint for motor acts depicting a closed fist whereby the object was not touched (8) was correctly categorized as a non-grasping action by both subjects. Finally, we tested videos depicting two novel hand configurations interacting with the object, during which an open hand either touched the object with the palm downwards (12) or upwards (13). Interestingly, both monkeys categorized the hand palm down videos correctly as a non-grasping motor act (12), but failed to categorize the hand palm-up motor act (13). Since performance was around chance level for the latter, neither monkey confused this action with a grasp, such as they apparently did for the closed fist touching the object (11). Rather, they presumably had problems attributing this non-familiar action to either of the 2 classes, as seemed to be the case for the artificial prosthetic arm in M2 (5).

## Discussion

Action recognition is an important aspect of social cognition. Successful interaction in cooperative or competitive situations requires the ability to infer the goals and intentions of others' actions. Especially since the discovery of mirror neurons<sup>2,31</sup>, there has been an increased interest in the neural substrates for action recognition, both in human and non-human primates.

Some of current theories as to how the brain processes information related to others' actions<sup>14,32–34</sup> are based to a large degree on single-cell evidence, describing selective neuronal responses during the observation of different actions and action settings. Although these studies have been instrumental in showing how different stimulus aspects and characteristics can influence neuronal responses, most have not directly tested action comprehension. Hence it remains difficult to conclude from such studies which aspects of the actions, such as the underlying goals and intentions of the actor or the efficacy of the action with respect to the environmental constraints, a monkey comprehends when observing these actions.

Various behavioral protocols, also used in developmental and comparative research in human infants and apes, have been tested to investigate monkeys' action comprehension abilities more directly with respect to understanding the goals and intentions of others' actions. Some of these studies investigated monkeys' behavioral responses while they observed rational versus irrational actions or accidental versus intentional actions<sup>8,10,12,35</sup>, while yet others assessed if monkeys can make a distinction between an actor that is unable versus one unwilling to perform a certain action<sup>9,13</sup>.

The main aim of our study was to use a two-alternative categorization task, compatible with simultaneous neurophysiological measurements, to examine if monkeys could discriminate different types of observed motor acts and were able to indicate in a straightforward manner their choice with a quantifiable behavioral response. Combining this type of operant behavioral experiments with current reversible perturbation techniques such as microstimulation, opto- and chemogenetics<sup>36–38</sup>, might provide additional insights into the crucial role of certain brain regions for action recognition and discrimination.

Our results show the feasibility of training monkeys to discriminate grasping motor acts from non-grasping types of manual motor acts, which are otherwise very similar. Although training took a substantial length of time (especially with monkey M1), both monkeys became very proficient in this task. Moreover, generalization tests suggest that monkeys did not merely memorize the videos or specific low-level details of the stimuli to solve the task, but could transfer the learned categorization rule to a wide range of novel untrained videos. Monkey M2, as opposed to monkey M1, had prior experience with performing a visually-guided saccade task<sup>39</sup> before learning the action categorization task, which might partly explain the faster learning curve of monkey M2. However, both monkeys reached overall similar levels in performance after around 10 sessions when only a few stimuli were introduced to the stimulus batch, indicating monkey M1 did not have an overall problem learning the saccade task. While the lengthy training in our study was partly due to the setup of the experiment (introducing a large number of different stimuli to the stimuli batch over a longer period of time) and our rather strict threshold of expected performance on the task (above 80%), similar behavioral studies show that overall task performance on these tasks are both related to task difficulty and individual differences between subjects. For instance, Vangeneugden *et al.*<sup>15</sup> trained monkeys on a similar task requiring their monkeys to discriminate actions depicting humanoid walkers. While the 3 monkey subjects in that study reached over 90% correct performance in as little as 10 training sessions when discrimination leftward versus rightward walking, individual differences in training length and overall performance became much more pronounced when monkeys had to discriminate forward versus backward walkers. While one of the subjects in that study reached near perfect performance on this task after 40 training sessions, their monkey M2 for instance never reached above 80% correct performance, even after 60 sessions.

Our behavioral study resulted in several interesting findings. Our data suggest that our monkeys did not interpret novel videos showing a mechanical grasping effector (prosthetic hand) in a manner similar to novel human or conspecific effectors performing the same grasp. While our experiment was not designed to test goal understanding per se, our finding seems in line with ethological observations made in rhesus monkeys and developmental studies in infants. Wood *et al.*<sup>35</sup> used a two-option social foraging context and examined whether macaques would comprehend tool-related (non-biological) actions as goal-directed. While monkeys were able to interpret human actions that belonged to their own motor repertoire as goal-directed, they failed to do so when observing actions done with a pole or machete. Likewise, developmental studies suggest that 6-month-old infants respond differently to grasping actions performed by a human or an artificial effector, suggesting that infants considered only human actions as goal-directed at that early age<sup>40</sup>.



Although monkeys failed to generalize to an artificial grasp, one cannot conclude that monkeys fail to acquire a concept of the functionality of artificial effectors or tools, but prior observational<sup>41</sup> or motor training<sup>42</sup> seems to be a requirement for achieving such generalization. Previously, we have shown that rhesus monkeys can learn about the functionality of tools and are able to acquire the highly skilled motor control necessary to operate these devices<sup>43</sup>. Our supplementary footage (Supplementary Video S6) shows that monkeys not only can use pliers to retrieve food in a controlled way, but they also understand the functionality of certain tool features, as evidenced by their ability to correctly rotate the tool to retrieve the food in an efficient manner. Future experiments will be needed to investigate the degree to which actions done with tools or non-biological effectors generalize to novel untrained non-biological actions, after monkeys have acquired experience, either through association learning (as in our categorization task) or through physical motor experience with these tools<sup>43,44</sup>.

While our experiment required monkeys to discriminate ‘grasping’ from ‘non-grasping’ videos, it is difficult to assess with certainty what visual aspects of the videos the monkeys used to solve the categorization task. Our generalization data suggest that monkeys did not solve the task by merely memorizing, nor by discriminating between motor acts that involve an effector interacting with an object versus effector movements that were directed away from the object. A finger or an open hand palm touching an object were correctly categorized as non-grasping, while a non-biological effector (prosthetic arm) interacting with an object failed to be categorized correctly. The fact that a closed fist touching an object was consistently discriminated as a grasp in both animals, suggests that particular combinations of local features, for example a thumb and additional fingers in close proximity to the object may have been a particularly informative cue with which to solve the categorization task.

Since monkeys were required to fixate the entire duration of the videos before making a saccade to a target, we cannot know during which phase of the observed actions monkeys made their decision. In order to investigate this, the task could be altered by showing the 2 peripheral targets together with the action video and by allowing the monkey to make a choice at any stage of the video presentation. Our current task also required monkeys to make a saccade to the same target location for a specific action class. In order to avoid a motor bias, which doesn’t affect the current behavioral data but which could be detrimental during electrophysiology or functional imaging experiments, the task could be altered by associating each action class with a target cue differing in shape or color. After video presentation, these targets could then be displayed randomly in a balanced fashion at different locations. As an alternative, instead of using saccades as a behavioral response measurement, monkeys could also be trained to use their hands to indicate their decision, either pressing down on a button or lifting their hand<sup>39</sup>. When combining this categorization task with a measurement of functional brain activity to gain insights in the neural correlates of action recognition, baseline trials could be added during which monkeys fixate only a central fixation point and afterwards make their motor response (saccade or hand movement) to select either of the 2 peripheral targets to receive their reward. In addition, in order to understand the specificity of potential causal effects of focal perturbations during these type of action discrimination tasks, it will be useful also to contrast these type of tasks with similar categorization tasks requiring judgements on non-action stimuli, or action versus non-action stimuli. Finally, we should mention that we only tested a limited number of generalization settings. Related to recent observations in humans<sup>18,45</sup>, it may be interesting to investigate in future experiments if discrimination of observed actions in monkeys is either dependent or independent of viewpoint.

Unfamiliarity with the observed motor acts seems to be an important factor for failure to generalize to novel action displays in our experiment. Both monkeys had problems interpreting the prosthetic arm grasping and the inverted hand touching an object. While it is difficult to disentangle whether this unfamiliarity primarily reflects perceptual or motor unfamiliarity (in the case of the inverted hand), a behavioral study that investigated monkeys’ ability to evaluate the effectiveness of goal-directed motor acts<sup>10</sup> showed that monkeys detect the efficacy of the goals of observed motor acts, provided they belong to the observer’s motor repertoire. In this study, monkeys’ capacity for understanding the efficacy of goal-directed and familiar motor acts failed to generalize to unfamiliar motor acts. On the other hand, other behavioral research in free-ranging rhesus monkeys suggests when assessing action outcomes, monkeys do not need to have exact motor representations of these actions<sup>46</sup>. With respect to visual familiarity, it is of interest that infant research shows that while 9-month-old infants do not interpret actions with a mechanical claw as goal-directed, perceptual exposure to a human operating that device, allowed 9-month old infants to interpret these actions as goal-directed<sup>47</sup>.

Our inverted-hand motor act that appeared in the generalization test (13) resembles a condition used in several behavioral studies investigating monkeys’ ability to differentiate intentional from accidental actions<sup>8,12,35</sup>. For instance, Wood and co-authors<sup>8,35</sup> used a visually similar action (hand flop) and reported that monkeys (tamarins and rhesus monkeys) as well as apes (chimpanzees), selected containers potentially containing food more often when a human experimenter would touch it with an intentional motor act (hand grasp), as compared to an accidental condition, in which the human experimenter dropped his hand in a reversed manner (palm facing upwards) onto the container. The authors suggest that because this motor act is not part of the monkeys natural motor repertoire, this may have led to an inability to comprehend this gesture as a goal-directed action<sup>35</sup>. Although they would be physically capable of producing this type of motor behavior, they presumably have no actual motor experience with it. Using similar actions, including an inverted-hand flop as Wood and co-authors<sup>35</sup>, Costes-Thiré and co-workers<sup>12</sup>, on the other hand, failed to find evidence for discrimination between accidental and intentional actions in Tonkean macaques and capuchin monkeys. While investigating intention recognition was beyond the scope of our current experiments, similar categorization tests such as the one described here might be informative for investigating whether monkeys can indeed discriminate accidental from intentional actions.

## Conclusions

Our study shows the feasibility of training monkeys on an active action observation task in which monkeys not only had to observe actions, but had to discriminate the observed motor acts and indicate their choices with a straightforward behavioral measure. Generalization testing showed a high degree of transfer to untrained

novel action videos. These types of tasks, in combination with neurophysiological recordings and reversible perturbations, will be instrumental in gaining a more in-depth understanding of the neural correlates of action recognition.

## References

- Jellema, T. & Perrett, D. I. Neural representations of perceived bodily actions using a categorical frame of reference. *Neuropsychologia*. **44**, 1535–46 (2006).
- Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. Action recognition in the premotor cortex. *Brain*. **119**, 593–609 (1996).
- Umiltà, M. A. *et al.* I Know What You Are Doing. *Neuron*. **31**, 155–165 (2001).
- Kohler, E. *et al.* Hearing sounds, understanding actions: action representation in mirror neurons. *Science*. **297**, 846–8 (2002).
- Fogassi, L. *et al.* Parietal lobe: from action organization to intention understanding. *Science*. **308**, 662–7 (2005).
- Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G. & Orban, G. A. Observing others: multiple action representation in the frontal lobe. *Science*. **310**, 332–6 (2005).
- Jastorff, J., Popivanov, I. D., Vogels, R., Vanduffel, W. & Orban, G. A. Integration of shape and motion cues in biological motion processing in the monkeys STS. *Neuroimage*. **60**, 911–21 (2012).
- Wood, J. N., Glynn, D. D., Phillips, B. C. & Hauser, M. D. The perception of rational, goal-directed action in nonhuman primates. *Science*. **317**, 1402–5 (2007).
- Phillips, W., Barnes, J. L., Mahajan, N., Yamaguchi, M. & Santos, L. R. ‘Unwilling’ versus ‘unable’: capuchin monkeys’ (Cebus apella) understanding of human intentional action. *Dev Sci*. **12**, 938–45 (2009).
- Rochat, M. J., Serra, E., Fadiga, L. & Gallese, V. The evolution of social cognition: goal familiarity shapes monkeys’ action understanding. *Curr Biol*. **18**, 227–32 (2008).
- Kupferberg, A., Glasauer, S. & Burkart, J. M. Do robots have goals? How agent cues influence action understanding in non-human primates. *Behav Brain Res*. **246**, 47–54 (2013).
- Costes-Thiré, M. *et al.* Evidence that monkeys (Macaca tonkeana and Sapajus apella) read moves, but no evidence that they read goals. *J Comp Psychol*. **129**, 304–10 (2015).
- Canteloup, C. & Meunier, H. ‘Unwilling’ versus ‘unable’: Tonkean macaques’ understanding of human goal-directed actions. *PeerJ*. **5**, e3227 (2017).
- Giese, M. A. & Rizzolatti, G. Neural and Computational Mechanisms of Action Processing: Interaction between Visual and Motor Representations. *Neuron*. **88**, 167–80 (2015).
- Vangeneugden, J., Vancleef, K., Jaeggli, T., VanGool, L. & Vogels, R. Discrimination of locomotion direction in impoverished displays of walkers by macaque monkeys. *J Vis*. **10**, 22.1–19 (2009).
- Nelissen, K. & Vanduffel, W. Goal understanding in non-human primates: active action categorization tasks. *J Vis*. **13**, 349 (2013).
- Bonini, L. The Extended Mirror Neuron Network: Anatomy, Origin, and Functions. *Neuroscientist*. **23**, 56–67 (2016).
- Platonov, A. & Orban, G. A. Action observation: the less-explored part of higher-order vision. *Sci Rep*. **6**, 36742 (2016).
- Vanduffel, W., Zhu, Q. & Orban, G. A. Monkey cortex through fMRI glasses. *Neuron*. **83**, 533–50 (2014).
- Salzman, C. D., Britten, K. H. & Newsome, W. T. Cortical microstimulation influences perceptual judgements of motion direction. *Nature*. **346**, 174–7 (1990).
- Afraz, A., Boyden, E. S. & DiCarlo, J. J. Optogenetic and pharmacological suppression of spatial clusters of face neurons reveal their causal role in face gender discrimination. *PNAS*. **112**, 6730–5 (2015).
- Verhoef, B. E., Vogels, R. & Janssen, P. Inferotemporal cortex subserves three-dimensional structure categorization. *Neuron*. **73**, 171–82 (2012).
- Moeller, S., Crapse, T., Chang, L. & Tsao, D. Y. The effect of face patch microstimulation on perception of faces and objects. *Nat Neurosci*. **20**, 743–752 (2017).
- Grossman, E. D., Battelli, L. & Pascual-Leone, A. Repetitive TMS over posterior STS disrupts perception of biological motion. *Vis Res*. **45**, 2847–53 (2005).
- Michael, J. *et al.* Continuous theta-burst stimulation demonstrates a causal role of premotor homunculus in action understanding. *Psychol Sci*. **25**, 963–72 (2014).
- Jacquet, P. O. & Avenanti, A. Perturbing the action observation network during perception and categorization of actions’ goals and grips: state-dependency and virtual lesion TMS effects. *Cereb Cortex*. **25**, 598–608 (2015).
- Avenanti, A., Candidi, M. & Urgesi, C. Vicarious motor activation during action perception: beyond correlational evidence. *Front Hum Neurosci*. **7**, 185 (2013).
- Vanduffel, W. *et al.* Visual motion processing investigated using contrast agent-enhanced fMRI in awake behaving monkeys. *Neuron*. **32**, 565–77 (2001).
- Nelissen, K., Fiave, P. A. & Vanduffel, W. Decoding Grasping Movements from the Parieto-Frontal Reaching Circuit in the Nonhuman Primate. *Cereb Cortex*. **18**(Feb), 1–15 (2017).
- Vogels, R. Categorization of complex visual images by rhesus monkeys. Part 1: behavioral study. *Eur J Neurosci*. **11**, 1223–38 (1999).
- di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. Understanding motor events: a neurophysiological study. *Exp Brain Res*. **91**, 176–80 (1992).
- Perrett, D. I. *et al.* Frameworks of analysis for the neural representation of animate objects and actions. *J Exp Biol*. **146**, 87–113 (1989).
- Fleischer, F., Caggiano, V., Their, P. & Giese, M. A. Physiologically inspired model for the visual recognition of transitive hand actions. *J Neurosci*. **33**, 6563–80 (2013).
- Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M. & Rozzi, S. Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol Rev*. **94**, 655–706 (2014).
- Wood, J. N., Glynn, D. D. & Hauser, M. D. Rhesus monkeys’ understanding of actions and goals. *Soc Neurosci*. **3**, 60–68 (2008).
- Ekstrom, L. B., Roelfsema, P. R., Arsenault, J. T., Bonmassar, G. & Vanduffel, W. Bottom-up dependent gating of frontal signals in early visual cortex. *Science*. **321**, 414–7 (2008).
- Gerits, A. *et al.* Optogenetically induced behavioral and functional network changes in primates. *Curr Biol*. **22**, 1722–6 (2012).
- Eldridge, M. A. *et al.* Chemogenetic disconnection of monkey orbitofrontal and rhinal cortex reversibly disrupts reward value. *Nat Neurosci*. **19**, 37–9 (2016).
- Premereur, E., Janssen, P. & Vanduffel, W. Effector specificity in macaque frontal and parietal cortex. *J Neurosci*. **35**, 3446–59 (2015).
- Woodward, A., Sommerville, J. & Guajardo, J. How Infants Make Sense of Intentional Action. In B. Malle, L. Moses & D. Baldwin (Eds.), *Intention and Intentionality* (pp 149–169). Cambridge, MA: MIT Press (2001).
- Ferrari, P. F., Rozzi, S. & Fogassi, L. Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *J Cogn Neurosci*. **17**, 212–26 (2005).
- Hihara, S., Obayashi, S., Tanaka, M. & Iriki, A. Rapid learning of sequential tool use by macaque monkeys. *Physiol Behav*. **78**, 427–34 (2003).
- Peeters, R. *et al.* The representation of tool use in humans and monkeys: common and uniquely human features. *J Neurosci*. **29**, 11523–39 (2009).
- Rochat, M. J. *et al.* Responses of mirror neurons in area F5 to hand and tool grasping observation. *Exp Brain Res*. **204**, 605–16 (2010).

45. de la Rosa, S., Mieskes, S., Bühlhoff, H. H. & Curio, C. View dependencies in the visual recognition of social interactions. *Front Psychol.* **4**, 752 (2013).
46. Wood, J. N., Glynn, D. D. & Hauser, M. D. The uniquely human capacity to throw evolved from a non-throwing primate: an evolutionary dissociation between action and perception. *Biol Lett.* **3**, 360–4 (2007).
47. Hofer, T., Hauf, P. & Aschersleben, G. Infant's perception of goal-directed actions performed by a mechanical device. *Infant Behav Dev.* **28**, 466–480 (2005).

## Acknowledgements

We thank C. Van Eupen for training the animals and performing the generalization tests, A. Coeman, W. Depuydt, M. De Paep, C. Fransen, A. Hermans, P. Kayenbergh, G. Meulemans, I. Puttemans and S. Verstraeten for technical assistance and Dr. S. Raiguel for comments on the manuscript. This work was supported by Hercules II funds, Fonds Wetenschappelijk Onderzoek Vlaanderen (G.0.622.08, G.0.593.09, G.0.007.12), KU Leuven BOF-ZAP Startfinanciering (10/14), KU Leuven Program Financing PFV/10/008 and the European Union's Horizon 2020 Framework Program for Research and Innovation under Grant Agreement No 720270 (Human Brain Project SGA1).

## Author Contributions

K.N. and W.V. designed the study. K.N. performed the experiments, analyzed the data and prepared all figures. K.N. wrote the main manuscript text with comments from W.V.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-15378-6>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017